

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
Please do not report the images to the
Image Problem Mailbox.

SON-040245
Another Reference

TRANSLATION

PUBLICATION OF UNEXAMINED JAPANESE PATENT APPLICATION

(1)Publication No.:

Hei.9-90995 (90995/1997)

(43)Publication Date:

April 4, 1997

(21) Application No.: Hei.7-249889 (249889-1995)

(22) Filing Date: September 27, 1995

(71) Applicant: NEC Corporation
(Assignee)

(72) Inventor(s): Kazunori OZAWA

(73) Attorney: Patent Attorney, Naoki Kyomoto

(54) Title of Invention: SPEECH CODING APPARATUS

(57) [ABSTRACT]

[SUBJECT] To provide a speech coding apparatus in which a good sound quality even when a bit rate is relatively low.

[SOLVING MEASUREMENT] In a sound source quantization circuit 350 of the speech coding apparatus, a sound source signal is represented by a combination of a plurality of pulses, wherein at least one pulse is represented by a pre-determined number of bits, and wherein at least one pulse width dependent of the pulse position is preset.

SON-043245

Another Reference

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

Not discussed in the specification.

特開平9-90995

(43) 公開日 平成9年(1997)4月4日

(51) IntCl. ⁹	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/14			G 1 0 L 9/14	G
	9/18			J
H 0 3 M 7/30		9382-5K	H 0 3 M 7/30	E
				B

審査請求 有 請求項の数12 OL (全 24 頁)

(21) 出願番号 特願平7-249889

(22) 出願日 平成7年(1995)9月27日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 小澤 一範

東京都港区芝五丁目7番1号 日本電気株式会社内

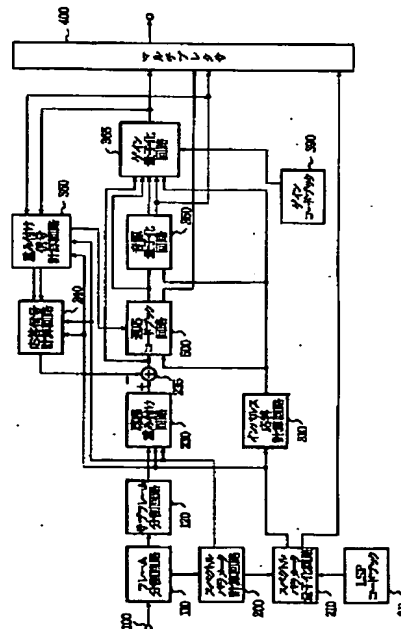
(74) 代理人 弁理士 京本 直樹 (外2名)

(54) 【発明の名称】 音声符号化装置

(57) 【要約】

【課題】 低ビットレートでも良好な音質の得られる音声符号化装置の提供。

【解決手段】 音声符号化装置の音源量子化回路350において、複数のパルスの組合せで音源信号を表わす。少なくとも一つのパルスをあらかじめ定められたビット数で表し、少なくとも一つのパルスの振幅が、パルスの位置に依存してあらかじめ決定されている。



【特許請求の範囲】

【請求項1】 入力した音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、少なくとも一つのパルスの位置をあらかじめ定められたビット数で表し、少なくとも一つのパルスの振幅がパルスの位置に依存しあらかじめ決定されている音源量子化部を有することを特徴とする音声符号化装置。

【請求項2】 少なくとも一つのパルスの振幅は、位置に依存してあらかじめ音声信号を用いて学習しておくことを特徴とする音源量子化部を有する請求項1記載の音声符号化装置。

【請求項3】 少なくとも一つのパルスのとりうる位置があらかじめ制限されている音源量子化部を有することを特徴とする請求項1または請求項2記載の音声符号化装置。

【請求項4】 入力した音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、少なくとも一つのパルスの位置をあらかじめ定められたビット数で表し、複数パルスの振幅をまとめて量子化する音源量子化部を有することを特徴とする音声符号化装置。

【請求項5】 複数パルスの振幅をまとめて量子化するために、あらかじめ音声信号を用いて学習して決定したコードブックを使用する音源量子化部を有することを特徴とする請求項4記載の音声符号化装置。

【請求項6】 少なくとも一つのパルスのとりうる位置があらかじめ制限されている音源量子化部を有することを特徴とする請求項4または請求項5記載の音声符号化装置。

【請求項7】 入力した音声信号からモードを判別し判別情報を出力するモード判定部と、前記音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記スペクトルパラメータを用いて音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、あらかじめ定められたモードの場合に、少なくとも一つのパルスの位置をあらかじめ定められたビット数で表し、少なくとも一つのパルスの振幅がパルスの位置に依存しあらかじめ決定されている音源量子化部を有することを特徴とする音声符号化装置。

【請求項8】 少なくとも一つのパルスの振幅は、位置に依存してあらかじめ音声信号を用いて学習して決定しておくことを特徴とする音源量子化部を有する請求項7記載の音声符号化装置。

【請求項9】 少なくとも一つのパルスのとりうる位置があらかじめ制限されている音源量子化部を有すること

を特徴とする請求項7または請求項8記載の音声符号化装置。

【請求項10】 入力した音声信号からモードを判別し判別情報を出力するモード判別部と、前記音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、あらかじめ定められたモードの場合に、少なくとも一つのパルスの位置をあらかじめ定められたビット数で表し、複数パルスの振幅をまとめて量子化する音源量子化部を有することを特徴とする音声符号化装置。

【請求項11】 複数パルスの振幅をまとめて量子化するために、あらかじめ音声信号を用いて学習して決定したコードブックを使用することを特徴とする音源量子化部を有する請求項10記載の音声符号化装置。

【請求項12】 少なくとも一つのパルスのとりうる位置があらかじめ制限されている音源量子化部を有することを特徴とする請求項10または請求項11記載の音声符号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声符号化装置に関し、特に、音声信号を低いビットレートで高品質に符号化する音声符号化装置に関する。

【0002】

【従来の技術】音声信号を高効率に符号化する方式としては、例えば、M.Schroeder and B.Atal氏による“Code-excited linear prediction: High quality speech at low bit rates”(Proc.ICASSP,pp.937-940,1985年)と題した論文(文献1)や、Kleijn氏による“Improved speech quality and efficeint vector quantization in SELP”(Proc.ICASSP,pp.155-158,1988年)と題した論文(文献2)などに記載されているCELP(Code Excited Linear Predictive Coding)が知られている。この従来例では、送信側では、フレームごと(例えば20ms)に音声信号から線形予測(LPC)分析を用いて、音声信号のスペクトル特性を表すスペクトルパラメータを抽出する。フレームにおけるサブフレーム(例えば5ms)に分割し、サブフレームごとに過去の音源信号を基に適応コードブックにおけるパラメータ(ピッチ周期に対応する遅延パラメータとゲインパラメータ)を抽出し、適応コードブックにより前記サブフレームの音声信号をピッチ予測する。ピッチ予測して求めた音源信号に対して、あらかじめ定められた種類の雑音信号からなる音源コードブック(ベクトル量子化コードブック)から最適な音源コードベクトルを選択し、最適なゲインを計算することにより、音源信号を量子化する。音源コードベクトルの選択の仕方は、選択した雑音信号により合成した信号と、前記残差信号との誤差電力を最小化する

ように行う。そして、選択されたコードベクトルの種類を表すインデックスとゲインならびに、前記スペクトルパラメータと適応コードブックのパラメータをマルチプレクサ部により組み合わせて伝送する。受信側の説明は省略する。

【0003】

【発明が解決しようとする課題】前記従来法では、音源コードブックから最適な音源コードベクトルを選択するのに多大な演算量を要するという問題がある。これは、文献1や2の方法では、音源コードベクトルを選択するのに、各コードベクトルに対して一旦フィルタリングもしくは畳み込み演算を行ない、この演算をコードブックに格納されているコードベクトルの個数だけ繰り返すことに起因する。例えば、コードブックのビット数がBビットで、次元数がNのときは、フィルタリングあるいは畳み込み演算のときのフィルタあるいはインパルス応答長をKとすると、演算量は1秒当たり、 $N \times K \times 2^B \times 8000 / N$ だけ必要となる。一例として、 $B=10$ 、 $N=40$ 、 $k=10$ とすると、1秒当たり81,920,000回の演算が必要となり、極めて膨大であるという問題点がある。

【0004】音源コードブック探索に必要な演算量を大幅に低減する方法として、種々のものが提案されている。例えば、ACELP (Algebraic Code Excited Linear Prediction) 方式が提案されている。これは、例えば、C.Laflammeらによる“16 kbps wideband speech coding technique based on algebraic CELP”と題した論文 (Proc. ICASSP, pp.13-16, 1991) (文献3)などを参照することができる。文献3の方法によれば、音源信号を複数のパルスで表し、各パルスの位置をあらかじめ定められたビット数で表し伝送する。ここで、各パルスの振幅は+1.0もしくは-1.0に限定されているため、振幅を伝送する必要はない。さらに、このために、パルス探索の演算量を大幅に低減化できる。

【0005】文献3の従来法では、演算量を大幅に低減化することが可能となるが、音質も充分ではないという問題点がある。この理由としては、各パルスが正負の極性のみか有しておらず、絶対値振幅はパルスの位置によらず常に1.0であるため、振幅を極めて粗く量子化したことになり、このために音質が劣化している。

【0006】本発明の目的は、上述の問題を解決し、ビットレートが低い場合でも、比較的少ない演算量で音質の劣化の少ない音声符号化装置を提供することにある。

【0007】

【課題を解決するための手段】本発明によれば、入力した音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、少なくとも一つのパルスの位置をあらかじめ定めら

れたビット数で表し、少なくとも一つのパルスの振幅がパルスの位置に依存しあらかじめ決定されている音源量子化部を有することを特徴とする音声符号化装置が得られる。

【0008】本発明によれば、少なくとも一つのパルスの振幅は、位置に依存してあらかじめ音声信号を用いて学習しておくことを特徴とする音源量子化部を有する請求項1記載の音声符号化装置が得られる。

【0009】本発明によれば、少なくとも一つのパルスのとりうる位置があらかじめ制限されている音源量子化部を有することを特徴とする請求項1または請求項2記載の音声符号化装置が得られる。

【0010】本発明によれば、入力した音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、少なくとも一つのパルスの位置をあらかじめ定められたビット数で表し、複数パルスの振幅をまとめて量子化する音源量子化部を有することを特徴とする音声符号化装置が得られる。

【0011】本発明によれば、複数パルスの振幅をまとめて量子化するために、あらかじめ音声信号を用いて学習して決定したコードブックを使用する音源量子化部を有することを特徴とする請求項4記載の音声符号化装置が得られる。

【0012】本発明によれば、少なくとも一つのパルスのとりうる位置があらかじめ制限されている音源量子化部を有することを特徴とする請求項4または請求項5記載の音声符号化装置が得られる。

【0013】本発明によれば、入力した音声信号からモードを判別し判別情報を出力するモード判定部と、前記音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記スペクトルパラメータを用いて音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、あらかじめ定められたモードの場合に、少なくとも一つのパルスの位置をあらかじめ定められたビット数で表し、少なくとも一つのパルスの振幅がパルスの位置に依存しあらかじめ決定されている音源量子化部を有することを特徴とする音声符号化装置が得られる。

【0014】本発明によれば、少なくとも一つのパルスの振幅は、位置に依存してあらかじめ音声信号を用いて学習して決定しておくことを特徴とする音源量子化部を有する請求項7記載の音声符号化装置が得られる。

【0015】本発明によれば、少なくとも一つのパルスのとりうる位置があらかじめ制限されている音源量子化部を有することを特徴とする請求項7または請求項8記載の音声符号化装置が得られる。

【0016】本発明によれば、入力した音声信号からモ

ードを判別し判別情報を出力するモード判別部と、前記音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、あらかじめ定められたモードの場合に、少なくとも一つのパルスの位置をあらかじめ定められたビット数で表し、複数パルスの振幅をまとめて量子化する音源量子化部を有することを特徴とする音声符号化装置が得られる。

【0017】本発明によれば、複数パルスの振幅をまとめて量子化するために、あらかじめ音声信号を用いて学習して決定したコードブックを使用することを特徴とする音源量子化部を有する請求項10記載の音声符号化装置が得られる。

【0018】本発明によれば、少なくとも一つのパルスのとりうる位置があらかじめ制限されている音源量子化部を有することを特徴とする請求項10または請求項11記載の音声符号化装置が得られる。

【0019】第1の発明では、音源量子化部において、一定時間間隔毎に、M個のパルスを立てて音源を量子化すると考える。i番目のパルスの振幅、位置をそれぞれ、 q_i 、 m_i とする。このとき、音源信号は下式のよう表せる。

【0020】

$$v(n) = G \sum_{i=1}^M g_i \delta(n - m_i) \quad (1)$$

【0021】ここで、Gは全体のレベルを表すゲインである。少なくとも一つのパルス、例えば、2つのパルスについて、位置の組合せの各々に対して、パルスの位置に依存して、あらかじめ振幅値を決定しておく。

【0022】第2の発明では、第1の発明におけるパルスの振幅値を、位置に依存して、あらかじめ、多量の音声信号を用いて学習して決定しておく。

【0023】第3の発明では、少なくとも一つのパルスのとりうる位置があらかじめ制限されている。例えば、偶数番目のサンプル位置、奇数番目のサンプル位置、Lサンプルとびのサンプル位置、などが考えられる。

【0024】第4の発明では、式(1)において、複数個のパルス(例えば2パルス)の振幅を表す振幅パターンをBビット分(2^B 種類)、振幅コードブックとしてあらかじめ用意しておき、最適な振幅パターンを選択する。

【0025】第5の発明では、第4の発明におけるBビット分の振幅コードブックをあらかじめ、多量の音声信号を用いて学習して決定しておく。

【0026】第6の発明では、第4または第5の発明において、少なくとも一つのパルスのとりうる位置があらかじめ制限されている。例えば、偶数番目のサンプル位

置、奇数番目のサンプル位置、Lサンプルとびのサンプル位置、などが考えられる。

【0027】第7の発明では、入力音声フレームごとに分割し、フレームごとに特徴量を使用してモードを判別する。以下ではモードの種類は4とする。モードは概ね次のように対応する。モード0：無音/子音部、モード1：過渡部、モード2：母音の弱定常部、モード3：母音の強定常部。そして、あらかじめ定められたモードの場合に、少なくとも一つのパルス、例えば、2つのパルスについて、位置の組合せの各々に対して、パルスの位置に依存してあらかじめ振幅値を決定しておく。

【0028】第8の発明では、第7の発明におけるパルスの振幅値をあらかじめ、多量の音声信号を用いて学習して決定しておく。

【0029】第9の発明では、第7または第8の発明において少なくとも一つのパルスのとりうる位置があらかじめ制限されている。例えば、偶数番目のサンプル位置、奇数番目のサンプル位置、Lサンプルとびのサンプル位置、などが考えられる。

【0030】第10の発明では、入力音声フレームごとに分割し、フレームごとに特徴量を使用してモードを判別する。そして、あらかじめ定められたモードの場合に、複数個のパルス(例えば2パルス)の振幅を表す振幅パターンをBビット分(2^B 種類)振幅コードブックとしてあらかじめ用意しておき、最適なパターンを選択する。

【0031】第11の発明では、第10の発明におけるBビット分の振幅コードブックをあらかじめ、多量の音声信号を用いて学習して決定しておく。

【0032】第12の発明では、第10または第11の発明において、少なくとも一つのパルスのとりうる位置があらかじめ制限されている。例えば偶数番目のサンプル位置、奇数番目のサンプル位置、Lサンプルとびのサンプル位置、などが考えられる。

【0033】

【発明の実施の形態】次に、本発明の実施の形態について図面を参照して説明する。

【0034】図1は本発明による音声符号化装置の第1の実施の形態を示すブロック図である。

【0035】図1を参照すると、入力端子100から音声信号を入力し、フレーム分割回路110では音声信号をフレーム(例えば10ms)ごとに分割し、サブフレーム分割回路120では、フレームの音声信号をフレームよりも短いサブフレーム(例えば2ms)に分割する。

【0036】スペクトルパラメータ計算回路200は、少なくとも一つのサブフレームの音声信号に対して、サブフレーム長よりも長い窓(例えば24ms)をかけて音声信号を切り出してスペクトルパラメータをあらかじめ定められた次数(例えば $P=10$ 次)計算する。ここでスペクトルパラメータの計算には、周知のLPC分析や、Bu

rg分析などを用いることができる。ここでは、Burg分析を用いることとする。Burg分析の詳細については、中溝著による“信号解析とシステム同定”と題した単行本

(コロナ社1988年刊)の82~87頁(文献4)などに記載されているので説明は略する。さらにスペクトルパラメータ計算部200は、Burg法により計算された線形予測係数 α_i ($i=1, \dots, 10$)を量子化や補間に適したLSPパラメータに変換する。ここで、線形予測係数からLSPへの変換は、菅村他による“線スペクトル対(LSP)音声分析合成方式による音声情報圧縮”と題した論文(電子通信学会誌、J64-A、pp.599-606、1981年)(文献5)を参照することができる。例えば、第2、4サブフレームでBurg法により求めた線形予測係数を、LSPパラメータに変換し、第1、3サブフレームのLSPを直線補間により求めて、第1、3サブフレームのLSPを逆変換して線形予測係数に戻し、第1-4サブフレームの線形予測係数 α_{il} ($i=1, \dots, 10, l=1, \dots, 5$)を聴感重み付け回路230に出力する。また、第4サブフレームのLSPをスペクトルパラメータ量子化回路210へ出力する。

【0037】スペクトルパラメータ量子化回路210は、あらかじめ定められたサブフレームのLSPパラメータを効率的に量子化し、下式の歪みを最小化する量子化値を出力する。

【0038】

$$D_j = \sum_i W(i) [LSP(i) - QLSP(i)]^2 \quad (2)$$

【0039】ここで、 $LSP(i)$ 、 $QLSP(i)_j$ 、 $W(i)$ はそれぞれ、量子化前の i 次目のLSP、量子化後の j 番目の結果、重み係数である。

【0040】以下では、量子化法として、ベクトル量子化を用いるものとし、第4サブフレームのLSPパラメータを量子化するものとする。LSPパラメータのベクトル量子化の手法は周知の手法を用いることができる。具体的方法は例えば、特開平4-171500号公報(文献6)や特開平4-363000号公報(文献7)や、特開平5-6199号公報(文献8)や、T.Nomura et al.,による

“LSP Coding VQ-SVQ With Interpolation in 4.075kbps M-LCELP Speech Coder”と対した論文(Proc. Mobile Multimedia Communications, pp.B.2.5, 1993)(文献

$$x_z(n) =$$

$$d(n) - \sum_{i=1}^{10} \alpha_i d(n-i) + \sum_{i=1}^{10} \alpha_i \tau^i y(n-i) + \sum_{i=1}^{10} \alpha_i' \tau^i x_z(n-i)$$

(3)

【0046】ただし、 $n-i \leq 0$ のときは

【0047】

9)などを参照できるのでここでは説明を略する。

【0041】また、スペクトルパラメータ量子化回路210は、第4サブフレームで量子化したLSPパラメータをもとに、第1~第4サブフレームのLSPパラメータを復元する。ここでは、現フレームの第4サブフレームの量子化LSPパラメータと1つ過去のフレームの第4サブフレームの量子化LSPを直線補間して、第1~第3サブフレームのLSPを復元する。ここで、量子化前のLSPと量子化後のLSPとの誤差電力を最小化するコードベクトルを1種類選択した後、直線補間により第1~第4のサブフレームのLSPを復元する。さらに性能を向上させるためには、前記誤差電力を最小化するコードベクトルを複数候補選択したのちに、各々の候補について、累積歪を評価し、累積歪を最小化する候補と補間USPの組を選択するようにすることができる。詳細は、例えば、特願平5-8737号明細書(文献10)を参照することができる。

【0042】以上により復元した第1~3サブフレームのLSPと第4サブフレームの量子化LSPをサブフレームごとに線形予測係数 α'_{il} ($i=1, \dots, 10, l=1, \dots, 5$)に変換し、インパルス応答計算回路310へ出力する。また、第4サブフレームの量子化LSPのコードベクトルを表すインデックスをマルチプレクサ400に出力する。

【0043】聴感重み付け回路230は、スペクトルパラメータ計算回路200から、各サブフレームごとに量子化前の線形予測係数 α'_{il} ($i=1, \dots, 10, l=1, \dots, 5$)を入力し、前記文献1にもとづき、サブフレームの音声信号に対して聴感重み付けを行い、聴感重み付け信号を出力する。

【0044】応答信号計算回路240は、スペクトルパラメータ計算回路200から、各サブフレームごとに線形予測係数 α_{il} を入力し、スペクトルパラメータ量子化回路210から、量子化、補間して復元した線形予測係数 α'_{il} をサブフレームごとに入力し、保存されているフィルタメモリの値を用いて、入力信号を零 $d(n)=0$ とした応答信号を1サブフレーム分計算し、減算回路235へ出力する。ここで応答信号 $x_z(n)$ を下式で表される。

【0045】

9

$$y(n-i)=p(N+(n-i)) \quad (4)$$

$$x_z(n-i)=s_w(N+(n-i)) \quad (5)$$

【0048】ここでNはサブフレーム長を示す。 γ は、聴感重み付け量を制御する重み係数であり、下記の式(7)と同一の値である。 $s_w(n)$ 、 $p(n)$ は、それぞれ、重み付け信号計算回路の出力信号、後述の式(7)における右辺第1項のフィルタの分母の項の出力信号をそれぞれ示す。

【0049】減算回路235は、下式により、聴感重み

$$H_w(z)=\frac{1-\sum_{i=1}^{10}a_1z^{-i}}{1-\sum_{i=1}^{10}a_1\gamma^iz^{-i}} \cdot \frac{1}{1-\sum_{i=1}^{10}a_1\gamma^iz^{-i}} \quad (7)$$

【0053】適応コードブック回路500は、ゲイン量子化回路365から過去の音源信号 $v(n)$ を、減算回路235から出力信号 $x'_w(n)$ を、インパルス応答計算回路310からインパルス応答 $h_w(n)$ を入力する。ピッチ

$$D_T=\sum_{n=0}^N \delta x_w^2(n)-[\sum_{n=0}^N x_w(n)y_w^2(n-T)]^2/[\sum_{n=0}^N y_w^2(n-T)] \quad (8)$$

【0055】ここで、

【0056】

$$y_w(n-T)=v(n-T)*h_w(n) \quad (9)$$

【0057】であり、記号*は畳み込み演算を表す。

【0058】

$$\beta=\sum_{n=0}^N \delta x_w(n)y_w(n-T)/\sum_{n=0}^N y_w^2(n-T) \quad (10)$$

【0059】ここで、女性音や、子供の声に対して、遅延の抽出精度を向上させるために、遅延を整数サンプルではなく、小数サンプル値で求めてもよい。具体的な方法は、例えば、P.Kroonによる、“Pitch predictors with high terminal resolution”と対した論文(Proc. ICASSP, pp.661-664, 1990年)(文献11)などを参照することができる。

【0060】さらに、適応コードブック回路500は、下式に従いピッチ予測を行ない、予測残差信号 $e_w(n)$ を音源量子化回路350へ出力する。

【0061】

$$e_w(n)=x_w(n)-\beta v(n-T)*h_w(n) \quad (11)$$

【0062】音源量子化回路350は、前述したように、M個のパルスをたてるとする。少なくとも一つのパルスの位置をあらかじめ定められたビット数で量子化し、位置を表すインデックスをマルチプレクサ400に出力する。パルスにおける位置の探索法は、一パルスずつ逐次的に探索する種々の方法が提案されており、例えば、K.Ozawa氏らによる“A study on pulse search algorithms for multipulse excited speech coder reali 50

10

付け信号から応答信号をサブフレーム分減算し、 $x'_w(n)$ を適応コードブック回路500へ出力する。

【0050】

$$x'_w(n)=x_w(n)-x_z(n) \quad (6)$$

【0051】インパルス応答計算回路310は、z変換が下式で表される聴感重み付けフィルタのインパルス応答 $h_w(n)$ をあらかじめ定められた点数Lだけ計算し、適応コードブック回路500、音源量子化回路350へ出力する。

【0052】

に対応する遅延Tを下式の歪みを最小化するように求め、遅延を表すインデックスをマルチプレクサ400に出力する。

【0054】

zation,”と題した論文(文献12)などを参照できるので、ここでは説明を省略する。また、これ以外でも前記文献3に記された方法や、後述の式(16)―(21)を記した方法などを用いることもできる。

【0063】このとき、少なくとも一つのパルスの振幅は、位置に依存してあらかじめ定まっている。

【0064】ここでは、一例としてM個のうちの2個のパルスの振幅がこれらの2個のパルスの位置の組合せに依存してあらかじめ定まっているとする。いま、第1パルス、第2パルスともに2種類の位置をとりえるとすると、これら2パルスの振幅の例としてはパルスの位置の組合せとしては(1,1)(1,2)(2,1)(2,2)があり、位置の組合せに対応して振幅としては、例えば、(1.0,1.0)(1.0,0.1)(0.1,1.0)(0.1,0.1)などが考えられる。振幅は位置の組合せに応じてあらかじめ定められているので、振幅を表すための情報を伝送する必要はない。

【0065】なお、2個以外のパルスは、簡略化のために、位置に依存せずにあらかじめ定められた振幅、例えば、1.0、-1.0など、をもたせることもできる。

【0066】振幅、位置の情報はゲイン量子化回路365に出力される。

【0067】ゲイン量子化回路365は、ゲインコードブック390からゲインコードベクトルを読みだし、選択された音源コードベクトルに対して、下式を最小化するようにゲインコードベクトルを選択する。ここでは、適応コードブックのゲインと音源のゲインの両者を同時にベクトル量子化する例について示す。

【0068】

11

12

$$D_k = \sum_{n=0}^{N-1} [x_w(n) - \beta'_k v(n-T) + h_w(n) - G'_k \sum_{i=1}^M g_i h_w(n-m_i)]^2 \quad (12)$$

【0069】ここで、 β'_k 、 G'_k は、ゲインコードブック390に格納された2次元ゲインコードブックにおけるk番目のコードベクトルである。選択されたゲインコードベクトルを表すインデックスをマルチプレクサ400に出力する。

【0070】重み付け信号計算回路360は、スペクトルパラメータ計算回路200の出力パラメータおよび、それぞれのインデックスを入力し、インデックスからそれに対応するコードベクトルを読みだし、まず下式にもとづき駆動音源信号 $v(n)$ を求める。

【0071】

$$s_w(n) = v(n) - \sum_{i=1}^{10} a_i v(n-i) + \sum_{i=1}^{10} a_i \gamma_i^{20} p(n-i) + \sum_{i=1}^{10} a'_i \gamma_i s_w(n-i) \quad (14)$$

【0075】図2は本発明の第2の実施の形態を示すブロック図である。この実施の形態は、図1の実施の形態に比して、音源量子化回路355の動作が異なる。ここでは、パルスの振幅値は、振幅パターンとして振幅パラメータ格納回路359に格納しておき、パルスの位置情報を入力して読み出す。このパターンは、パルスの位置の組合せに依存して、多量の音声データベースを用いて学習し、位置に依存して一意に決定しておく。

【0076】図3は本発明の第3の実施の形態を示すブロック図である。音源量子化回路357では、各パルスのとりうる位置があらかじめ制限されている。例えば、偶数番目のサンプル位置、奇数番目のサンプル位置、Lサンプルとびのサンプル位置、などが考えられる。ここでは、サンプルとびのサンプル位置をとることにし、Lの値は次のように選ぶ。

【0077】

$$L = N/M$$

(15)

$$D_k = \sum_{n=0}^{N-1} [e_w(n) - G \sum_{i=1}^M g'_{ik} h_w(n-m_i)]^2 \quad (16)$$

【0082】ここで、Gは最適ゲイン、 g'_{ik} は、k番目の振幅コードベクトルにおけるi番目のパルス振幅である。

【0083】式(16)の最小化は以下のように定式化

$$v(n) = \beta'_k v(n-T) + G'_k \sum_{i=1}^M \delta(n-m_i) \quad (13)$$

【0072】 $v(n)$ は適応コードブック回路500に出力される。

【0073】次に、重み付け信号計算回路360は、スペクトルパラメータ計算回路200の出力パラメータ、スペクトルパラメータ量子化回路210の出力パラメータを用いて下式により、応答信号 $s_w(n)$ をサブフレームごとに計算し、応答信号計算回路240に出力する。

【0074】

ここで、N、Mはそれぞれ、サブフレーム長、パルスの個数を示す。

【0078】なお、少なくとも一つのパルスの振幅は、パルスの位置に依存してあらかじめ決定されていてもよい。

【0079】図4は本発明の第4の実施の形態を示すブロック図である。音源量子化回路450は、第1の実施の形態と同一の方法でパルスの位置を求め、これを量子化してマルチプレクサ400およびゲイン量子化回路365へ出力する。

【0080】さらに、複数パルスの振幅をまとめてベクトル量子化する。具体的に説明すると、パルス振幅コードブック451から、パルス振幅コードベクトルを読みだし、下式の歪みを最小化する振幅コードベクトルを選択する。

【0081】

できる。式(16)をパルスの振幅 g'_i で偏微分して0とおくと

【0084】

$$D_k = \sum_{n=0}^{N-1} [e_w(n)/G]^2 - \left[\sum_{n=0}^{N-1} \phi(n) v_k(n) \right]^2 / \left[\sum_{i=1}^M g'_{ik} h_w(n-m_i) \right]^2 \quad (17)$$

【0085】ここで

【0086】

$$\phi(n) = \sum_{i=n}^{N-1} e_w(i) h_w(i-n), n=0, \dots, N-1 \quad (18)$$

$$v_k(n) = \sum_{i=1}^M g'_{ik} \delta(n-m_i) \quad (19)$$

$$\left[\sum_{i=1}^M g'_{ik} h_w(n-m_i) \right]^2 = \sum_{i=1}^M g'^2_{ik} \Psi(m_i, m_i) + 2 \sum_{i=1}^{M-1} \sum_{j=i+1}^M g'_{ik} g'_{jk} \Psi(m_i, m_j) \quad (20)$$

【0091】ここで

【0092】

$$\Psi(m_i, m_j) = \sum_{n=j}^{N-1} h(n-i) h(n-j), j \geq i \quad (21)$$

【0093】したがって、式(20)の $g'_{ik}{}^2$ と $g'_{ik} g'_{jk}$ を振幅コードベクトル k ごとにあらかじめ計算してコードブックに格納しておくことにより、計算量を大幅に低減化できる。また、サブフレームごとに ϕ と ψ を一度計算しておけば、さらに演算量を低減化できる。

【0094】この場合の振幅量子化に必要な積和回数は、サブフレーム当たりのパルスの個数を M とし、サブフレーム長を L 、インパルス応答長を L 、振幅コードブックのビット数を B とすると、サブフレーム当たり、概ね $N^2 + [(M-1)! + M] 2^B + NL + M 2^B$ となる。 $B=10$ 、 $N=40$ 、 $M=4$ 、 $L=20$ とすると、この値は、1秒当たり、3,347,200回となる。また、パルスの位置を探索するには、文献12に記載されている方式1を使用すれば、上記演算量に対して新たに発生する演算量はないので、文献1、2の従来方式の方法に比べ、約1/24

【0087】である。

【0088】したがって、式(16)の最小化は、式(17)の右辺第2項の最大化と等価となる。

【0089】式(17)の右辺第2項の分母は下式のように変形できる。

【0090】

20 となる。

【0095】したがって、本方法を用いることにより、パルスの振幅、位置探索に必要な演算量は、従来方式に比べ、極めて少ないことがわかる。

【0096】音源量子化回路は以上の方法で選択された振幅コードベクトルのインデックスをマルチプレクサ400に出力する。また、各パルスの位置と振幅コードベクトルによる各パルスの振幅をゲイン量子化回路365に出力する。

【0097】図5は図4の実施の形態の変形を示すブロック図である。音源・ゲイン量子化回路550では、ゲインを量子化しながらパルスの振幅の量子化を行なう点が、図4の音源量子化回路450と異なる。パルスの位置は音源量子化回路450と同一の方法で求め、同一の方法で量子化する。パルスの振幅とゲインは、下式を最小化するように、パルス振幅コードブック451、ゲインコードブック390からそれぞれ、パルス振幅コードベクトルとゲインコードベクトルを選択することにより、量子化する。

【0098】

$$D_k = \sum_{n=0}^{N-1} [x_w(n) - \beta'_k v(n-T) * h_w(n) - G'_k \sum_{i=1}^M g'_{ik} h_w(n-m_i)]^2 \quad (22)$$

【0099】ここで、 g'_{ik} は、 k 番目のパルス振幅コードベクトルにおける i 番目のパルス振幅である。

β'_k 、 G'_k は、ゲインコードブック390に格納された2次元ゲインコードブックにおける k 番目のコードベクトルである。パルス振幅ベクトルとゲインコードベクトルのすべての組合せに対し、式(22)を最小化するよ

うに最適な組合せを1組選択することができる。

【0100】また、探索演算量を低減化するために予測選択を導入することもできる。例えば、式(16)あるいは式(17)の歪みが小さい順にパルス振幅コードベクトルを複数個予備選択し、各候補に対してゲインコードブックを探索し、式(22)を最小化するパルス振幅

コードベクトルとゲインコードベクトルの組合せを1種類選択する。

【0101】選択されたパルス振幅コードベクトル、ゲインコードベクトルを表すインデックスをマルチプレクサ400に出力する。

【0102】図6は本発明の第5の実施の形態を示すブロック図である、図4の実施の形態に比して、パルス振幅学習コードブック580が異なる。このコードブックは、複数パルスの振幅を量子化するためのコードブックを、音声信号を用いてあらかじめ学習して格納しておく。コードブックの学習法は、例えば、Linde氏による“An algorithm for vector quantization design,”と題した論文（IEEE Trans. Commun., pp.84-95, January, 1980）（文献13）などを参照できる。

【0103】なお、図5と同様に、ゲインをゲインコードブックにより量子化しながら、パルス振幅をパルス振幅コードブックにより量子化するような構成にすることもできる。

【0104】図7は本発明の第6の実施の形態を示すブロック図である。図4の実施の形態に比して、音源量子化回路470が異なる。各パルスのとりうる位置からあらかじめ制限されている。例えば、偶数番目のサンプル位

$$P_i = \sum_{n=0}^{N-1} x_{wi}^2(n) \quad (24)$$

$$E_i = P_i - [\sum_{n=0}^{N-1} x_{wi}(n)x_{wi}(n-T)]^2 / [\sum_{n=0}^{N-1} x_{wi}^2(n-T)] \quad (25)$$

【0110】ここで、Tは予測ゲインを最大化する最適遅延である。

【0111】フレーム平均ピッチ予測ゲインGをあらかじめ複数個のしきい値と比較して複数種類のモードに分類する。モードの個数としては、例えば4を用いることができる。モード判別回路800は、モード判別情報を音源量子化回路600、マルチプレクサ400へ出力する。

【0112】音源量子化回路600は、モード判別情報があらかじめ定められたモードを示す場合に以下の処理を行なう。

【0113】式(1)に示すようにM個のパルスを求めるとし、少なくとも一つのパルスの位置をあらかじめ定められたビット数で量子化し、位置を表すインデックスをマルチプレクサに出力する。このとき、少なくとも一つのパルスの振幅は、位置に依存してあらかじめ定まっている。

【0114】ここでは、一例としてM個のうちの2個のパルスの振幅がこれらの2個のパルスの位置の組合せに依存してあらかじめ定まっているとする。いま、第1パルス、第2パルスともに2種類の位置をとりえらるとすると、これら2パルスの振幅の例としてはパルスの位置の組合せとしては(1,1)(1,2)(2,1)(2,2)があり、位置の組合せに対応して振幅としては、例えば、(1.0,1.0)(1.0,

置、奇数番目のサンプル位置、Lサンプルとびのサンプル位置、などが考えられる。ここでは、Lサンプルとびのサンプル位置をとることにし、Lの値は式(13)に示したように選ぶ。

【0105】なお、複数パルスの振幅をまとめてコードブックを用いて量子化することもできる。

【0106】図8は本発明の第7の実施の形態を示すブロック図である。モード判別回路800は、聴感重み付け回路230からフレーム単位で聴感重み付け信号を受取り、モード判別情報を出力する。ここでは、モード判別に、現在のフレームの特徴量を用いる。特徴量としては、例えば、フレームで平均したピッチ予測ゲインを用いる。ピッチ予測ゲインの計算は、例えば下式を用いる。

【0107】

$$G = 10 \log_{10} [1/L \sum_{i=1}^L (P_i/E_i)] \quad (23)$$

【0108】ここで、Lはフレームに含まれるサブフレームの個数である。P_i、E_iはそれぞれ、i番目のサブフレームでの音声パワ、ピッチ予測誤差パワを示す。

【0109】

0.1)(0.1,1.0)(0.1,0.1)などが考えられる。振幅は位置の組合せに応じてあらかじめ定められているので、振幅を表すための情報を伝送する必要はない。

【0115】なお、2個以上のパルスは、簡略化のために、位置に依存せずにあらかじめ定められた振幅、例えば、1.0、-1.0など、をもたせることもできる。

【0116】振幅、位置の情報はゲイン量子化回路365に出力される。

【0117】図9は本発明の第8の実施の形態を示すブロック図である。音源量子化回路650は、モード判別回路800から判別情報を入力し、あらかじめ定められたモードの場合に、振幅パラメータ格納回路359から、パルスの位置情報を入力して、パルスの振幅値を読み出す。

【0118】このパターンは、パルスの位置の組合せに依存して、多量の音声データベースを用いて学習し、位置に依存して一意に決定しておく。学習法については、前記文献13などを参照できる。

【0119】図10は本発明の第9の実施の形態を示すブロック図である。音源量子化回路680は、モード判別回路800から判別情報を入力し、あらかじめ定められたモードの場合に、各パルスのとりうる位置があらかじめ制限されている。例えば、偶数番目のサンプル位置、奇数番目のサンプル位置、Lサンプルとびのサン

ル位置、などが考えられる。ここでは、Lサンプルとびのサンプル位置をとることにし、Lの値は式(15)のように選ぶ。

【0120】なお、少なくとも一つのパルスの振幅を位置に依存してあらかじめ振幅パターンとして学習しておいてもよい。

【0121】図11は本発明の第10の実施の形態を示すブロック図である。音源量子化回路700は、モード判別回路800から判別情報を入力し、あらかじめ定められたモードの場合に、少なくとも一つのパルスの位置をあらかじめ定められたビット数で量子化し、インデックスをゲイン量子化回路365、マルチプレクサ400へ出力する。次に、複数パルスの振幅をまとめてベクトル量子化する。パルス振幅コードブック451から、パルス振幅コードベクトルを読みだし、式(14)の歪み最小化する振幅コードベクトルを選択する。そして、選択された振幅コードベクトルのインデックスをゲイン量子化回路365、マルチプレクサ400へ出力する。

【0122】なお、式(17)を用いて、ゲインを量子化しながら、パルス振幅を量子化する構成をとることもできる。

【0123】図12は本発明の第11の実施の形態を示すブロック図である。音源量子化回路750は、モード判別回路800から判別情報を入力し、あらかじめ定められたモードの場合に、少なくとも一つのパルスの位置をあらかじめ定められたビット数で量子化し、インデックスをゲイン量子化回路365、マルチプレクサ400へ出力する。次に、複数パルスの振幅をまとめてベクトル量子化する。パルス振幅学習コードブック580から、あらかじめ学習されたパルス振幅コードベクトルを読みだし、式(14)の歪み最小化する振幅コードベクトルを選択する。そして、選択された振幅コードベクトルのインデックスをゲイン量子化回路365、マルチプレクサ400へ出力する。

【0124】なお、式(22)を用いて、ゲインを量子化しながら、パルス振幅を量子化する構成をとることもできる。

【0125】図13は本発明の第12の実施の形態を示すブロック図である。音源量子化回路780は、モード判別回路800から判別情報を入力し、あらかじめ定められたモードの場合に、少なくとも一つのパルスの位置をあらかじめ定められたビット数で量子化する。ここで、各パルスのとりうる位置があらかじめ制限されている。例えば、偶数番目のサンプル位置、奇数番目のサンプル位置、Lサンプルとびのサンプル位置、などが考えられる。ここでは、Lサンプルとびのサンプル位置をとることにし、Lの値は式(15)のように選ぶ。インデックスをゲイン量子化回路365、マルチプレクサ400へ出力する。

【0126】なお、パルス振幅コードブックとしては、

第11の実施の形態で述べたように、あらかじめ学習したコードブックを使用することもできる。

【0127】さらに、式(22)を用いて、ゲインを量子化しながら、パルス振幅を量子化する構成をとることもできる。

【0128】上述した実施の形態に限らず、種々の変形が可能である。

【0129】モード判別情報を用いて適応コードブック回路や、ゲインコードブックを切替える構成とすることもできる。

【0130】

【発明の効果】以上説明したように、本発明によれば、音源量子化部において、少なくとも一つのパルスの位置はあらかじめ定められたビット数で量子化し、パルスの少なくとも1つのパルスの振幅がパルスの位置に依存してあらかじめ決定されているので、あるいは、パルスの振幅がパルスの位置に依存して、音声信号を用いてあらかじめ学習されているので、音源探索の演算量を低くおさえながら、従来方式よりも音質が改善される。

【0131】さらに、本発明によれば、複数パルスの振幅をまとめて量子化するためにコードブックを有しているので、音源探索の演算量を低くおさえながら、従来方式よりも音質が一層改善されるという大きな効果がある。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態のブロック図である。

【図2】本発明の第2の実施の形態のブロック図である。

【図3】本発明の第3の実施の形態のブロック図である。

【図4】本発明の第4の実施の形態のブロック図である。

【図5】図4の実施の形態の変形を示すブロック図である。

【図6】本発明の第5の実施の形態のブロック図である。

【図7】本発明の第6の実施の形態のブロック図である。

【図8】本発明の第7の実施の形態のブロック図である。

【図9】本発明の第8の実施の形態のブロック図である。

【図10】本発明の第9の実施の形態のブロック図である。

【図11】本発明の第10の実施の形態のブロック図である。

【図12】本発明の第11の実施の形態のブロック図である。

【図13】本発明の第12の実施の形態のブロック図である。

ある。

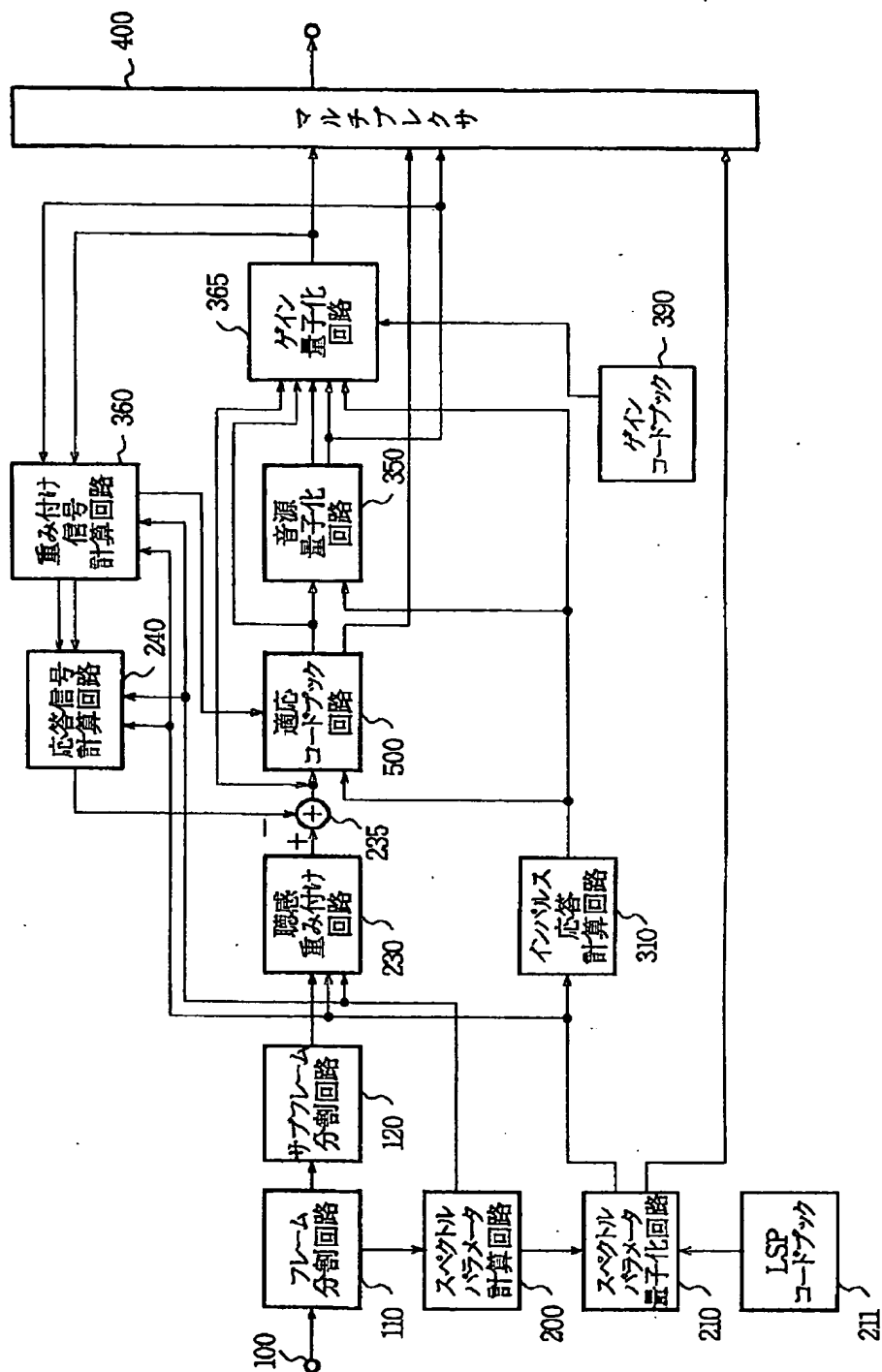
【符号の説明】

110 フレーム分割回路
 120 LSPパラメータ分割回路
 200 スペクトルパラメータ計算回路
 210 スペクトルパラメータ量子化回路
 211 LSPコードブック
 230 聴感重み付け回路
 235 減算回路
 240 応答信号計算回路
 310 インパルス応答計算回路
 350, 355, 357, 450, 470, 600, 650, 680, 700, 750, 780 音源量子化

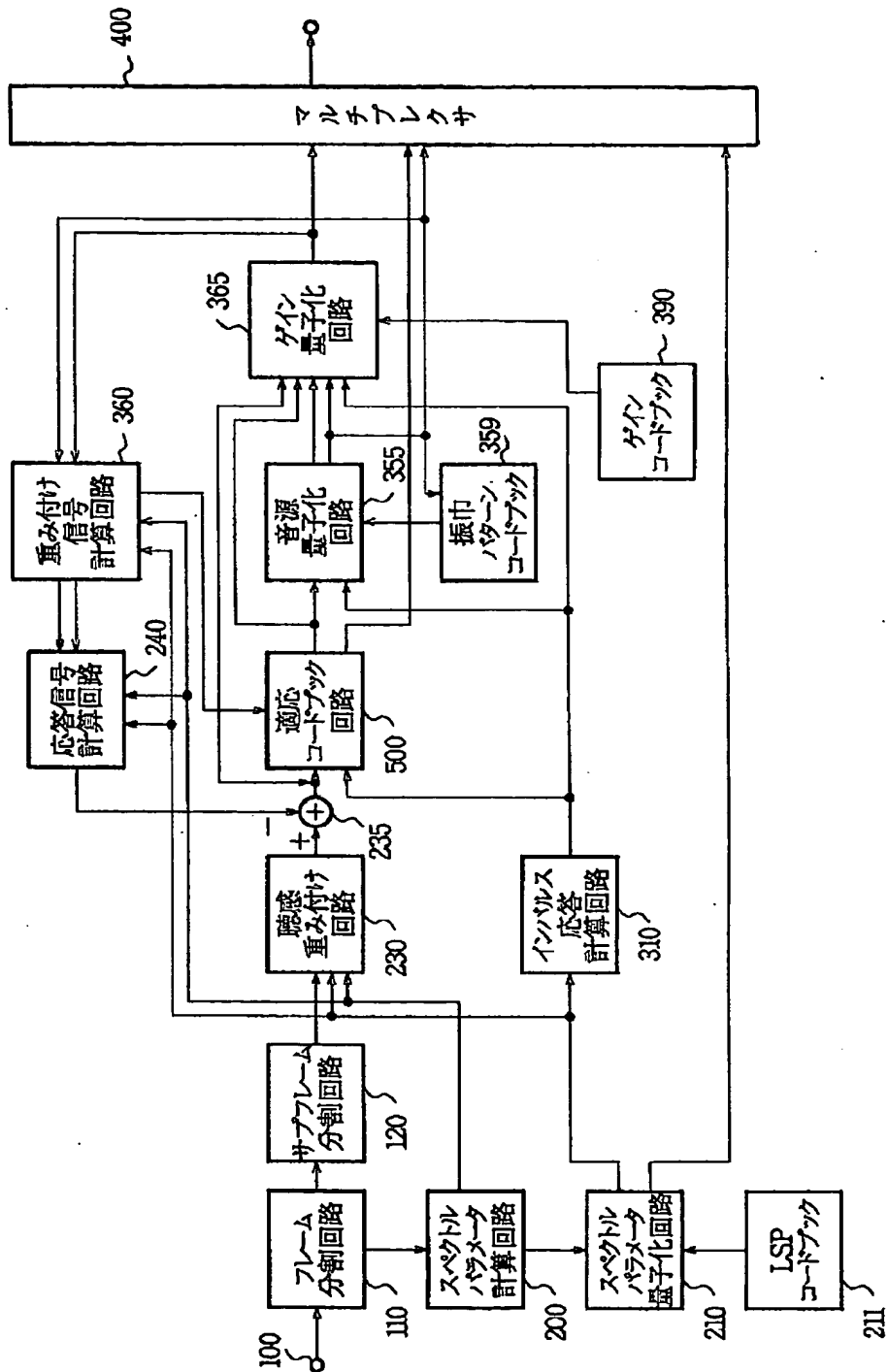
回路

500 適応コードブック回路
 550 音源・ゲイン量子化回路
 359 振幅パラメータ格納回路
 355 音源量子化回路
 360 重み付け信号計算回路
 365 ゲイン量子化回路
 390 ゲインコードブック
 400 マルチプレクサ
 10 451 パルス振幅コードブック
 580 パルス振幅学習コードブック
 800 モード判別回路

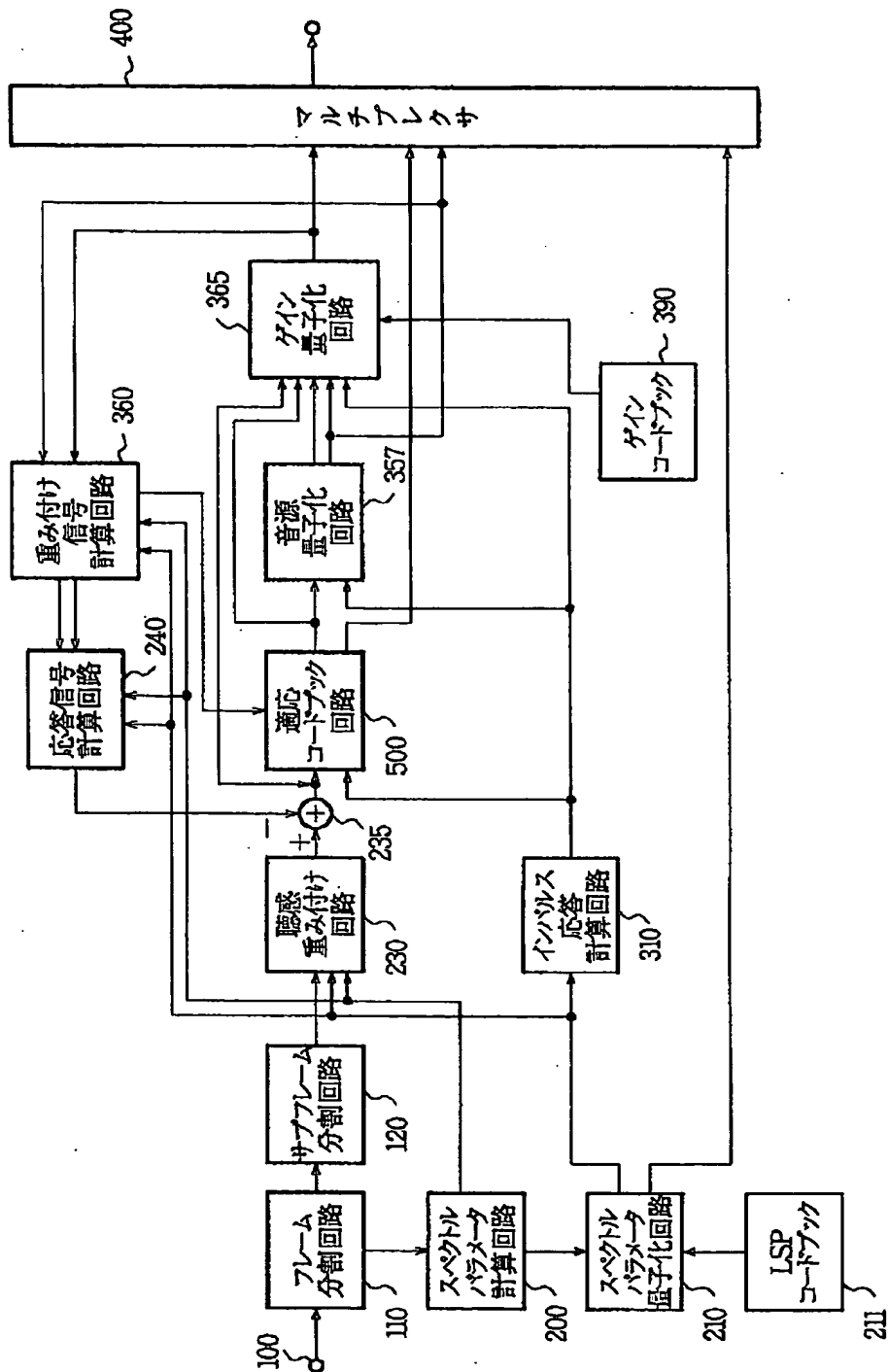
【図1】



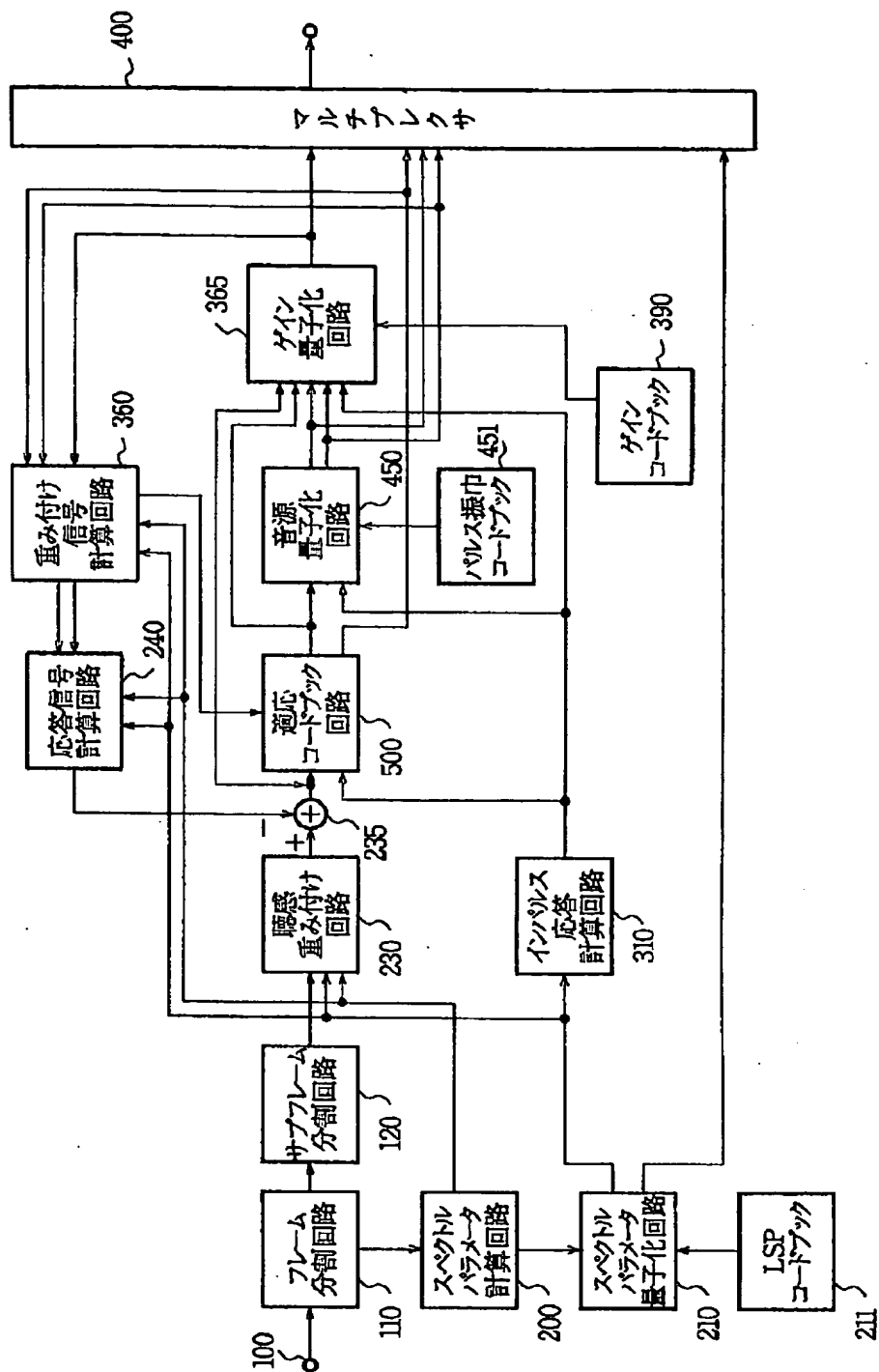
【図 2】



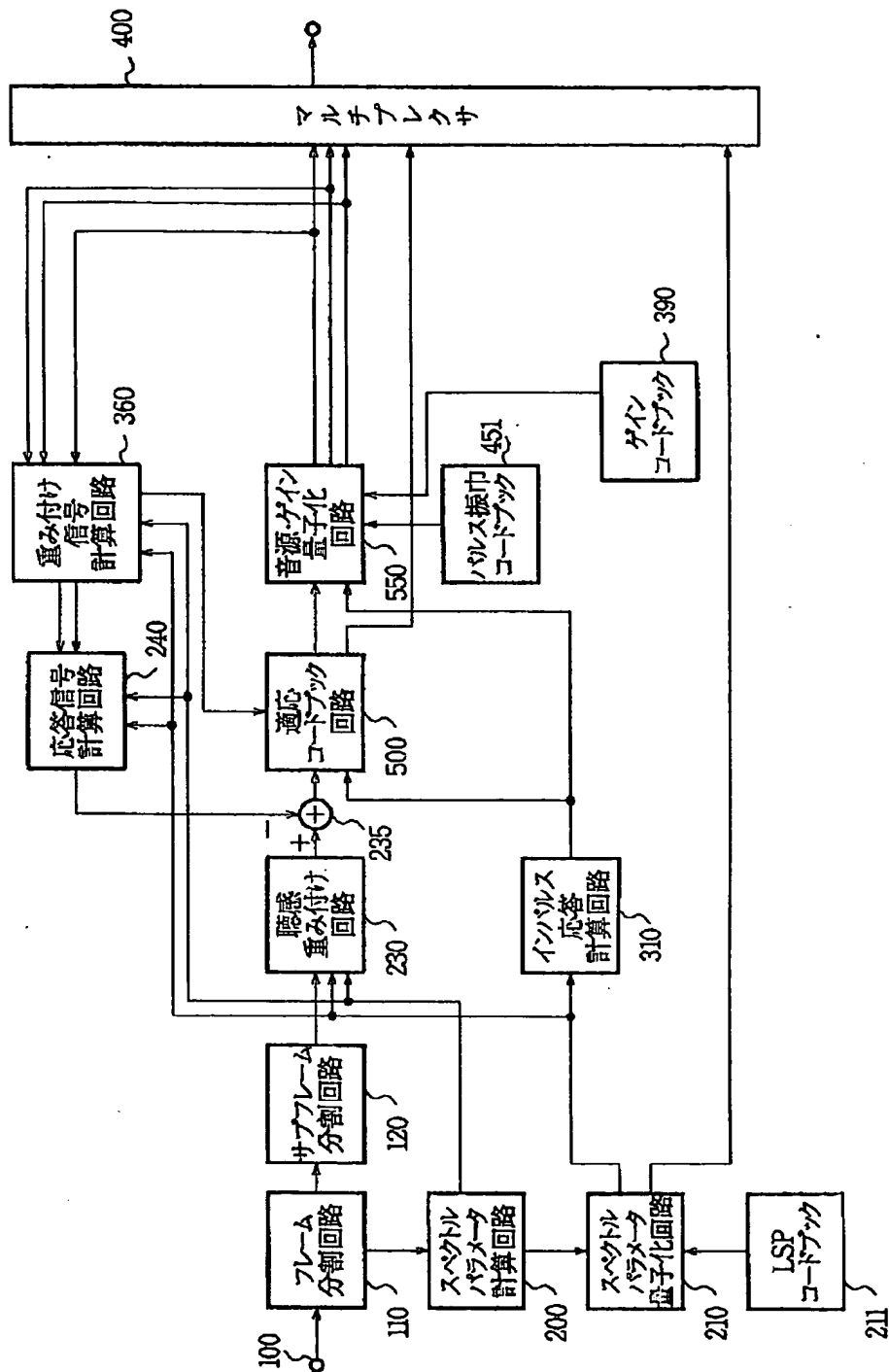
【例 3】



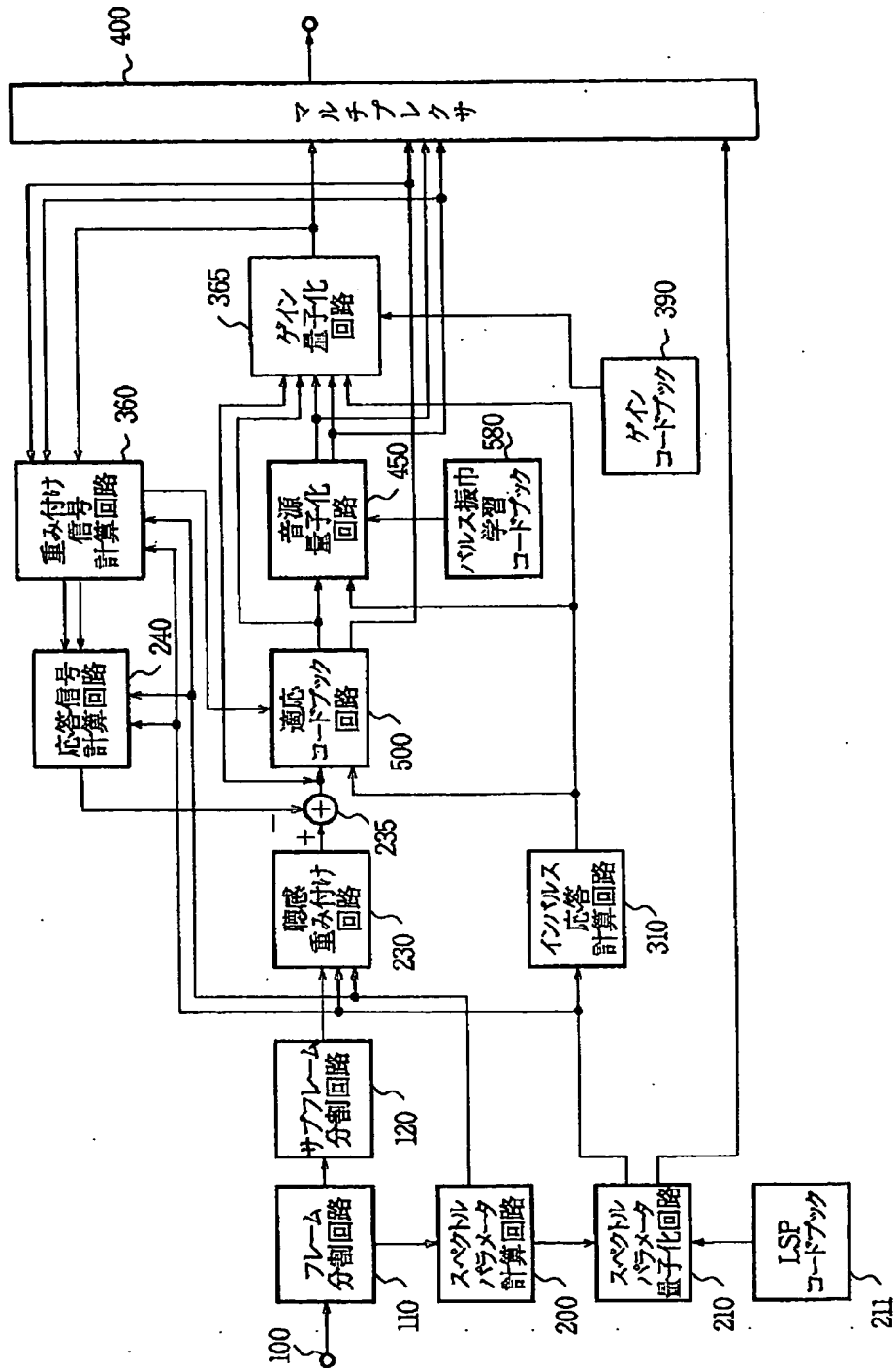
【図4】



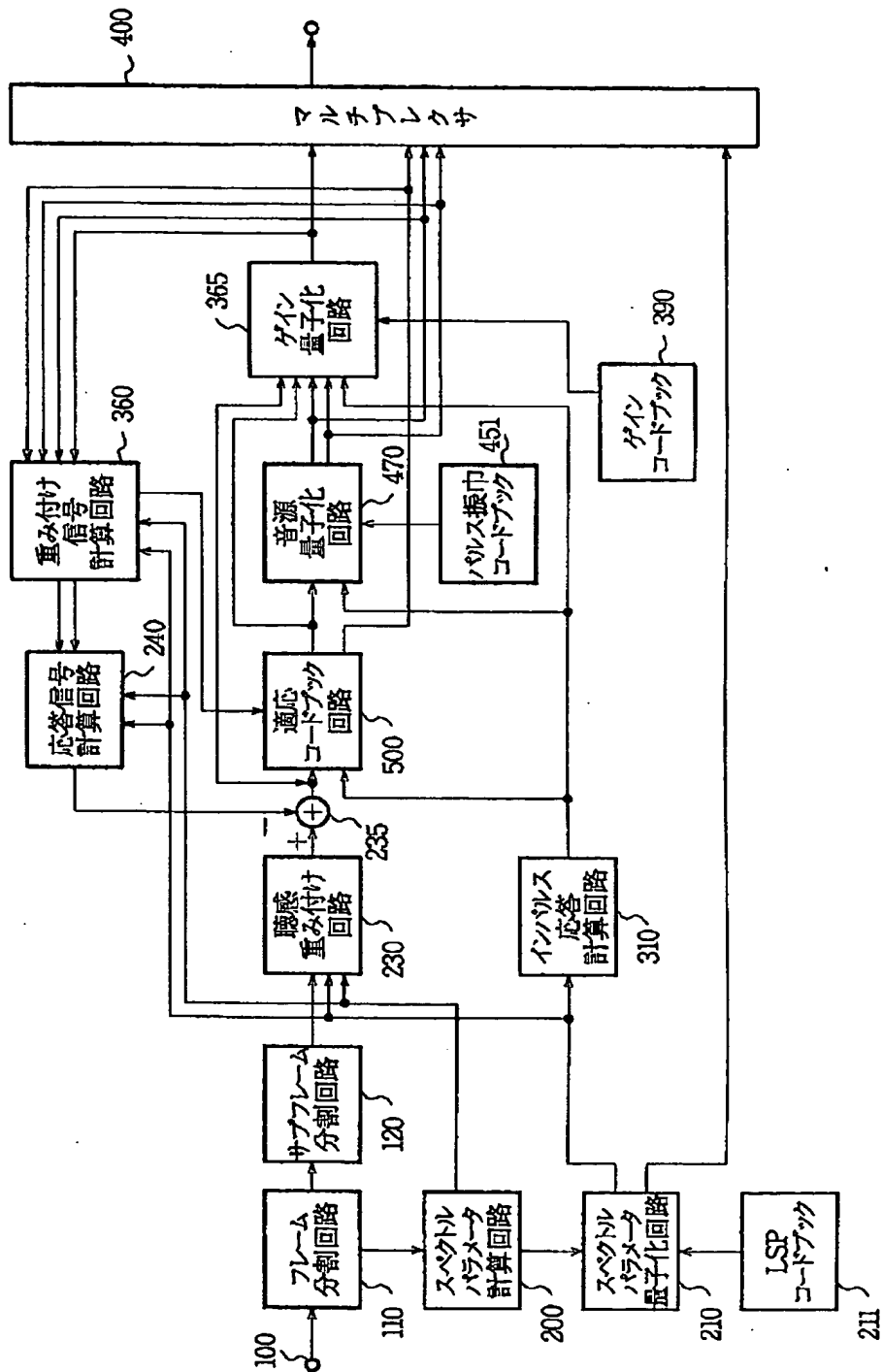
【図5】



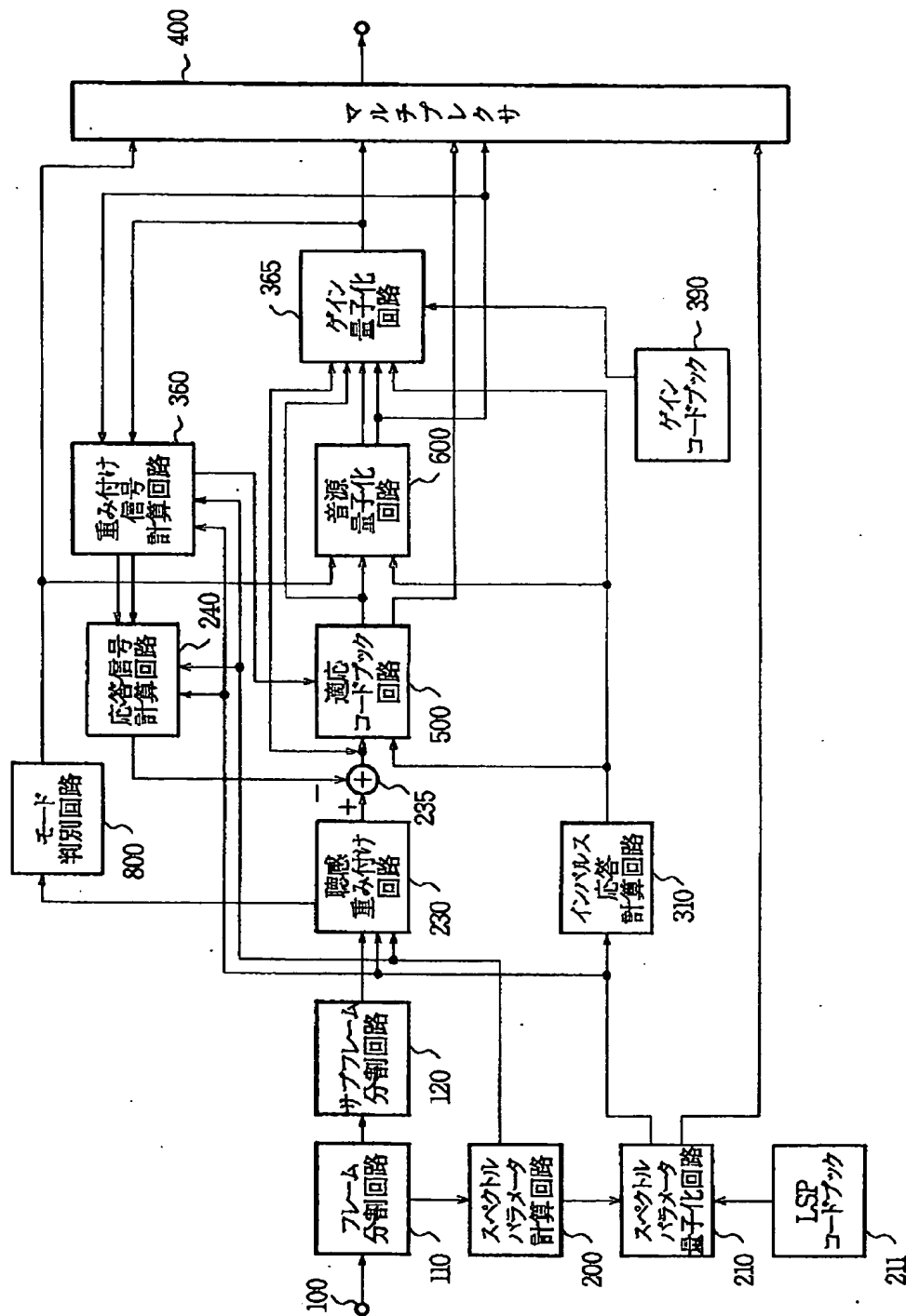
【図6】



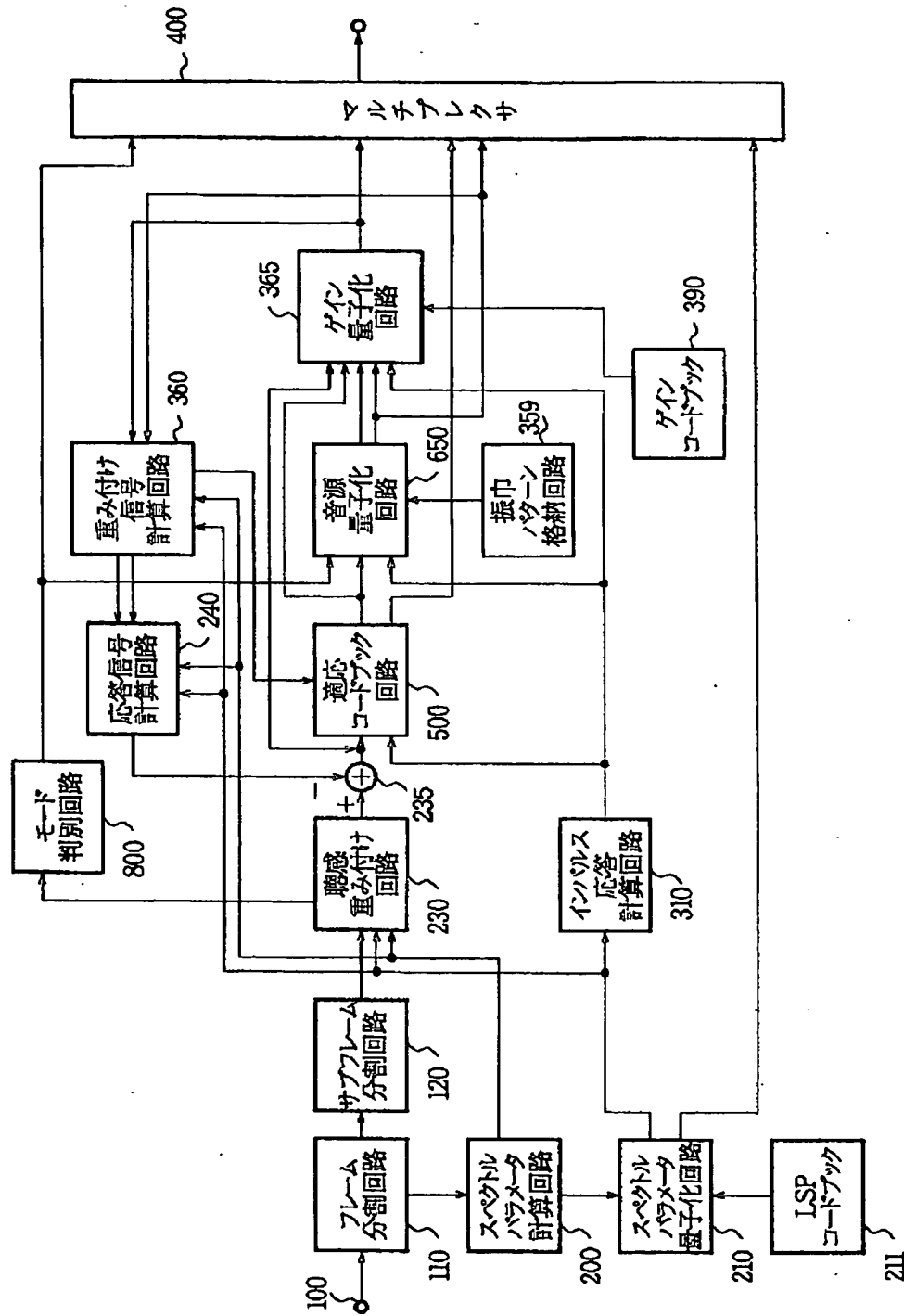
【図7】



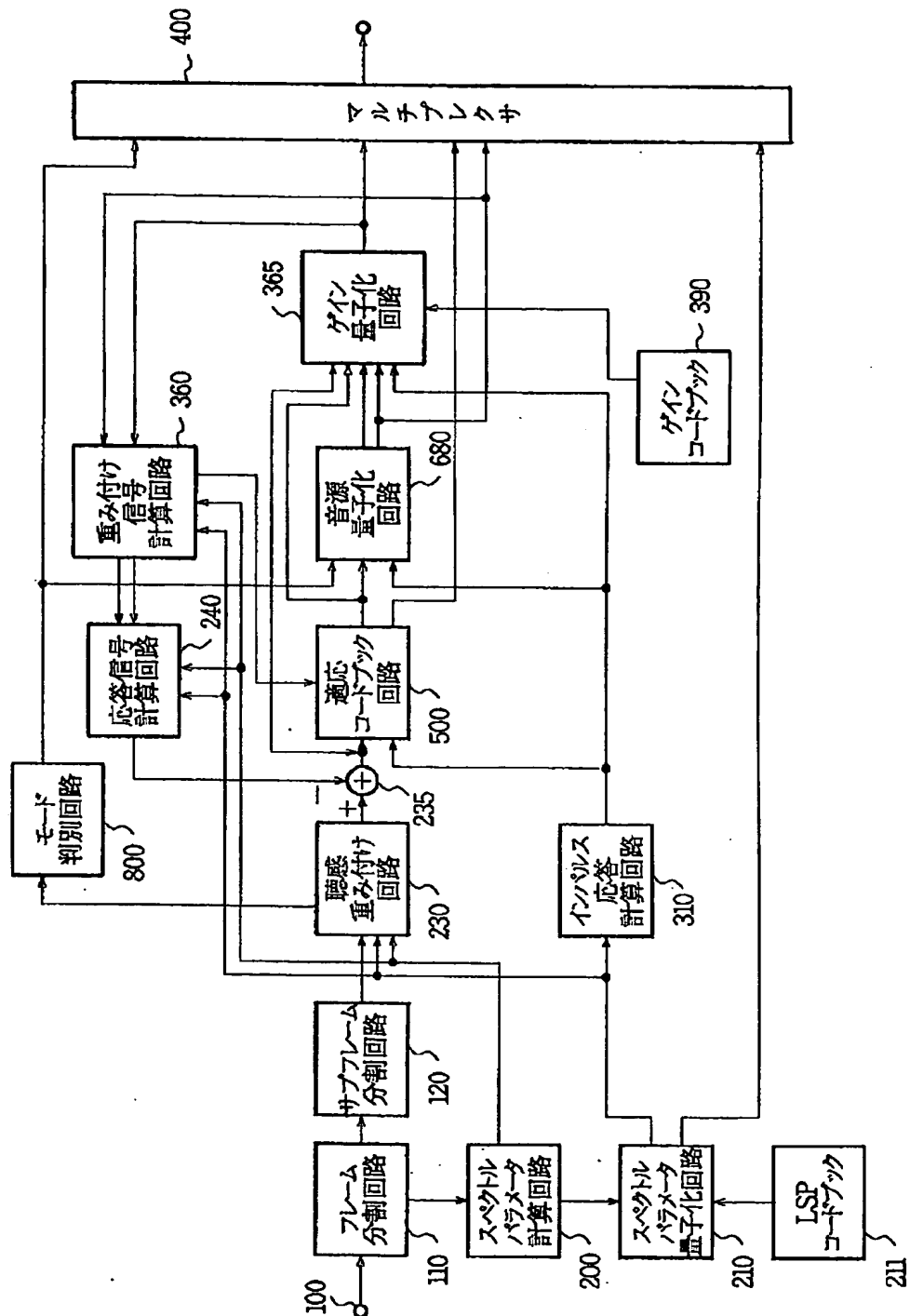
【図8】



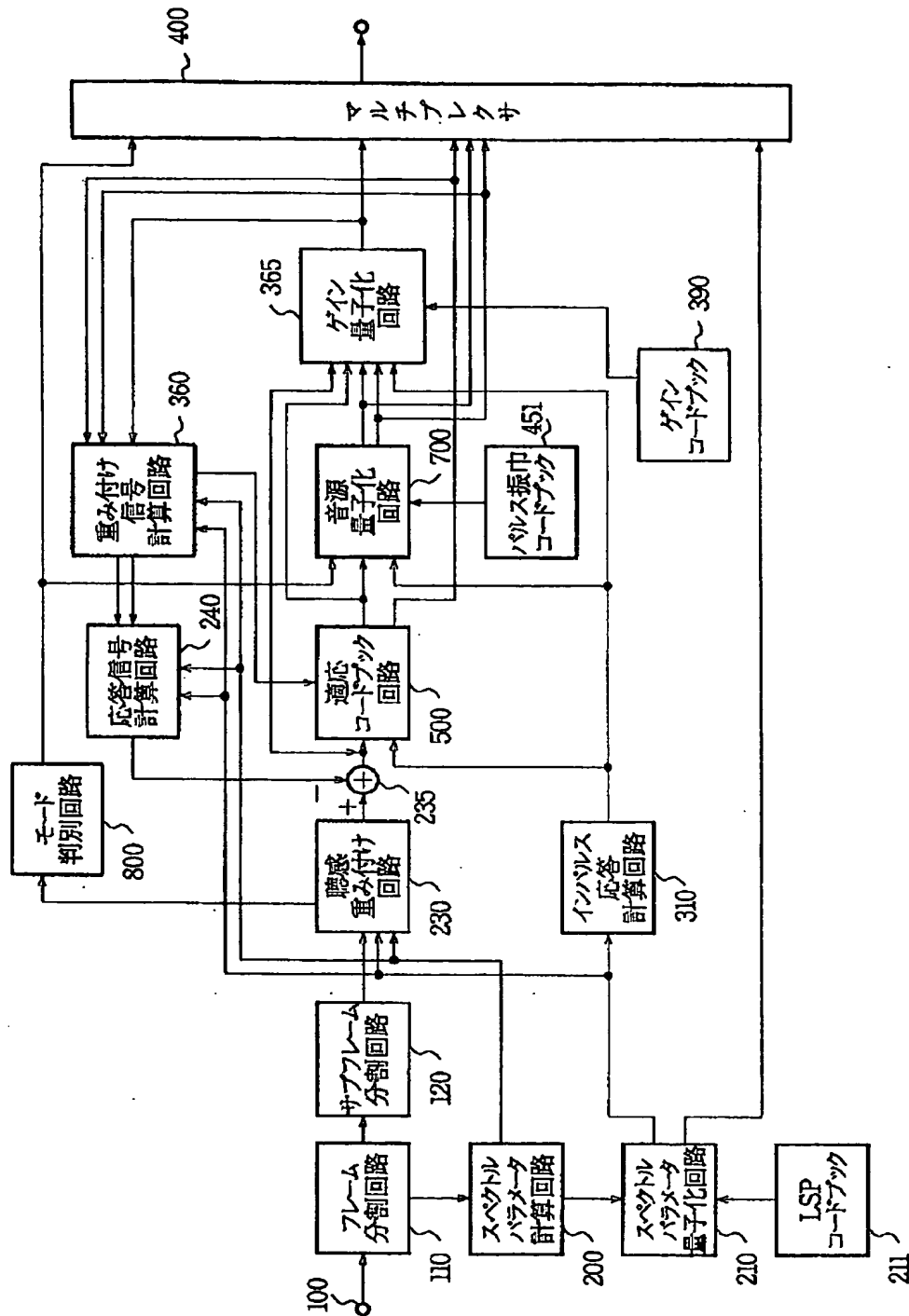
【図9】



【☒ 1 0】



【図11】



[illegible]